

AD-A036 517

TRAINING ANALYSIS AND EVALUATION GROUP (NAVY) ORLANDO FLA F/G 5/9  
TRAINING EFFECTIVENESS ASSESSMENT. VOLUME I. CURRENT MILITARY T--ETC(U)  
DEC 76 E R HALL, K LAM, S G BELLAMY  
TAE6-39-VOL-1

UNCLASSIFIED

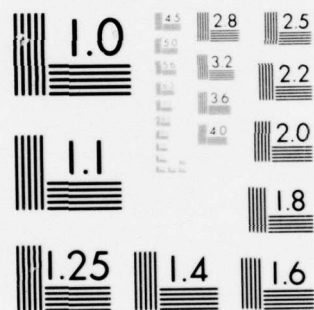
NL

| OF |  
AD  
A036517



END

DATE  
FILMED  
3-77



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

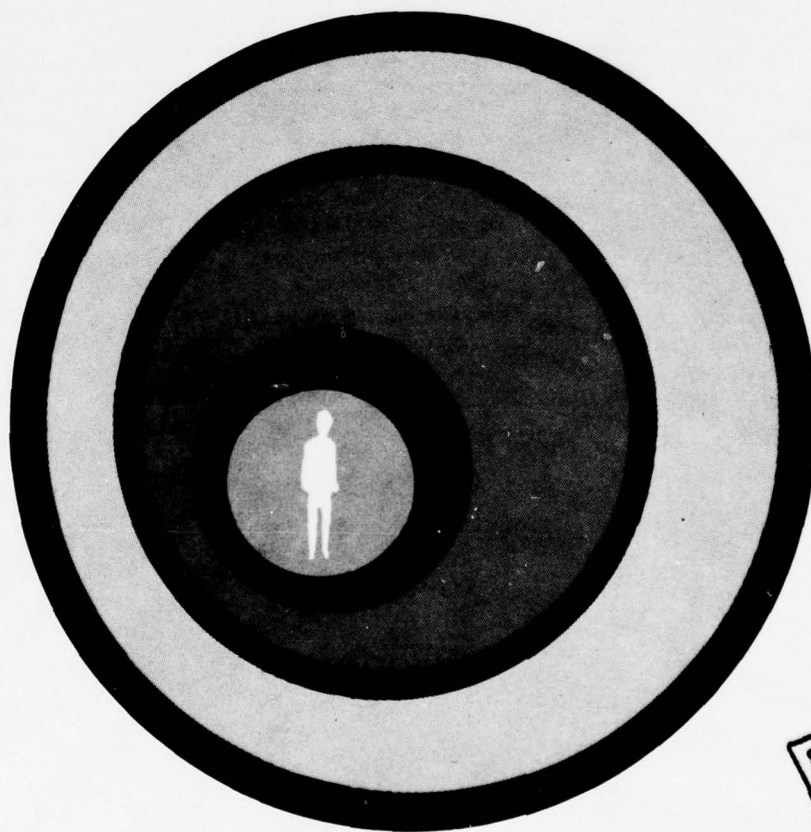
ADA036517

# T A E G

TRAINING  
ANALYSIS  
AND  
EVALUATION  
GROUP

EG REPORT  
39

13  
TRAINING EFFECTIVENESS ASSESSMENT:  
VOLUME I, CURRENT MILITARY TRAINING  
EVALUATION PROGRAMS



FOCUS  
ON  
THE  
TRAINED  
MAN

APPROVED FOR PUBLIC RELEASE;  
DISTRIBUTION IS UNLIMITED



DECEMBER 1976



TRAINING ANALYSIS AND EVALUATION GROUP  
ORLANDO, FLORIDA 32813

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14 TAEG Report No. - 39 - Vol - 1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 TRAINING EFFECTIVENESS ASSESSMENT: VOLUME I. CURRENT MILITARY TRAINING EVALUATION PROGRAMS.	5. TYPE OF REPORT & PERIOD COVERED 9 Final Report, May 1975 - Nov 1976.	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Eugene R. Hall, Karen Lam, and Susan G. Bellomy	8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Training Analysis and Evaluation Group Orlando, Florida 32813	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE 11 December 1976	13. NUMBER OF PAGES 40 12544p
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution is unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES  See also <u>Training Effectiveness Assessment: Volume II, Problems, Concepts, and Evaluation Alternatives.</u>		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Training                      Evaluation Training Evaluation      Military Training Training Effectiveness Assessment		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  A study was conducted to clarify issues and problems involved in the assessment of the effectiveness of military training and to evaluate and recommend objective procedures for determining the effectiveness of Navy training. The study results are reported in two volumes. This volume, volume I, reviews current military training evaluation programs. Evaluation philosophy, documentation and current practices in the assessment of training effectiveness within the United States Air Force, Navy, Marine Corps, and Army are described.		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE  
S/N 0102-LF-014-6601

Unclassified 407626  
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)



Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block #20 (continued)

Information is provided concerning strengths and apparent deficiencies of the programs which were in effect between June 1975 and May 1976. Volume II, Problems, Concepts, and Evaluation Alternatives, examines specific problems affecting the conduct of evaluation programs within the Navy. It also provides guidance for the conduct of training effectiveness assessments and describes and evaluates a variety of objective techniques suitable for this purpose.

ACCESSION FOR	
NIS	White Section <input checked="" type="checkbox"/>
BDC	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY GROUP	
Dist.	AVAIL and/or SPECIAL
A	

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TAEG Report No. 39

TRAINING EFFECTIVENESS ASSESSMENT: VOLUME I,  
CURRENT MILITARY TRAINING EVALUATION PROGRAMS

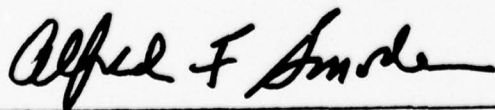
Eugene R. Hall  
Karen Lam  
Susan G. Bellomy

Training Analysis and Evaluation Group

December 1976

GOVERNMENT RIGHTS IN DATA STATEMENT

Reproduction of this publication in whole  
or in part is permitted for any purpose  
of the United States Government.



ALFRED F. SMODE, Ph.D., Director,  
Training Analysis and Evaluation Group



WORTH SCANLAND, Ph.D.  
Assistant Chief of Staff  
Research and Program Development  
Chief of Naval Education and Training

TAEG Report No. 39

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
I	INTRODUCTION. . . . .	5
	Purpose . . . . .	5
	Approach. . . . .	6
II	CURRENT MILITARY TRAINING EVALUATION PROGRAMS . . . . .	7
	U.S. Air Force Training Evaluation. . . . .	7
	Purpose of Program . . . . .	7
	Evaluation Responsibilities. . . . .	7
	Graduate Field Evaluation Program. . . . .	9
	Other Checks on Graduate Quality . . . . .	10
	Training Quality Reports. . . . .	10
	Inspector General Inspections . . . . .	11
	Job Performance Evaluations . . . . .	11
	Course Review Program. . . . .	11
	Student Critique Program . . . . .	12
	U.S. Navy Training Evaluation . . . . .	12
	CNET Instruction 1540.3, <u>Appraisal and</u> <u>Improvement of Training</u> . . . . .	13
	CNET Instruction 1540.3A (Proposed). . . . .	14
	CNET Instruction 1540.6, <u>Training Appraisal</u> <u>Subsystem</u> . . . . .	14
	CNTECHTRA ATO Manual . . . . .	15
	Internal Evaluation . . . . .	16
	Procedures . . . . .	16
	Utilization of Data. . . . .	17
	External Evaluation . . . . .	17
	Procedures . . . . .	17
	ISD Evaluation Program/Concepts. . . . .	17
	Internal Evaluation . . . . .	18
	External Evaluation . . . . .	18

TAEG Report No. 39

TABLE OF CONTENTS (continued)

<u>Section</u>	<u>Page</u>
Personnel Training and Evaluation Program (PTEP) . . . . .	19
U.S. Marine Corps Training Evaluation Programs. . . . .	20
Academic Regulations . . . . .	21
Evaluation and Validation . . . . .	21
Evaluation Documents. . . . .	22
Tests. . . . .	22
Instructional Rating Forms . . . . .	22
After Instruction Reports. . . . .	23
Graduate Questionnaires. . . . .	23
Design of Courses of Instruction . . . . .	23
Internal Evaluation . . . . .	24
Validation. . . . .	24
Feedback Evaluation. . . . .	24
Marine Corps Command and Staff College Evaluation Program. . . . .	25
Marine Corps Unit Level Training Management. . . . .	25
Marine Corps General Training System (GENTRAS) . . . . .	26
U.S. Army Training Evaluation . . . . .	26
Army Training and Evaluation Program (ARTEP) . . . . .	27
Methods . . . . .	28
Individual Evaluation. . . . .	28
III SUMMARY AND DISCUSSION. . . . .	31
Summary . . . . .	32
Discussion of Evaluation Programs . . . . .	32
U.S. Air Force . . . . .	32
Procedures. . . . .	32
Personnel . . . . .	33
Resources . . . . .	33

TAEG Report No. 39

TABLE OF CONTENTS (continued)

<u>Section</u>	<u>Page</u>
U.S. Navy . . . . .	33
Training Appraisal Plans. . . . .	33
Curriculum and Instructional Standards Offices. . . . .	35
CNTECHTRA A10 Manual. . . . .	36
U.S. Marine Corps . . . . .	36
U.S. Army . . . . .	36
REFERENCES . . . . .	38



## SECTION I

### INTRODUCTION

Each military service attempts to determine if the training it conducts is effective. Typically, these assessment efforts are designed to obtain both measures of the effectiveness of specific training courses and diagnostic information useful for improving those courses. There is considerable variability among the services in the way in which training evaluation is viewed and practiced. The unique viewpoints and procedures are of general interest and a review of these programs can provide useful insights as well as specific information for developing improved programs for evaluating Navy training and maintaining its quality.

The Chief of Naval Education and Training (CNET) tasked the Training Analysis and Evaluation Group (TAEG) to develop an assessment capability for determining the effectiveness of Navy training. Emphasis was to be on the identification and development of means for conducting assessments of training effectiveness.

#### PURPOSE

The overall study conducted in response to the tasking was concerned with organizing information relevant to the assessment of training effectiveness within a military setting. It was also concerned with the development of assessment methods suitable for use within the Navy environment. This effort is a necessary prelude to the subsequent development of systematized and standardized procedures for assessing training effectiveness that can be applied on a programmatic basis within the Navy.

The results of the study are reported in two volumes. This volume (I) examines and assesses the interservice practices and issues in training effectiveness assessment. Current practices of the United States Air Force, Navy, Marine Corps, and Army in the assessment of training effectiveness are reviewed to:

- . Identify and describe training evaluation efforts of the military services
- . Assess these efforts to identify the strengths, weaknesses, problem areas, constraints, etc., which characterize training evaluation in the military.

The intent is to reveal through examination of a representative sample of military evaluation programs the extent of efforts to control training quality and the quality or value of these efforts. This review addresses the current status and stature of training evaluation within the military.



## TAEG Report No. 39

A second, and perhaps more important, purpose is to glean from the concepts, principles, practices, and philosophies extant in training effectiveness assessment those aspects which might fruitfully be applied to the development of systems for assessing the effectiveness of Navy training.

Volume II, Problems, Concepts, and Evaluation Alternatives, utilizes information from the review. It examines the particular problems involved for determining training effectiveness within the Navy, provides guidance for conducting training effectiveness assessments, and evaluates various methodological approaches that are suitable for this purpose.

### APPROACH

For the review, much diverse documentation (e.g., manuals, directives, regulations, pamphlets, instructions) was obtained from the military services. These documents were critically examined to determine how each service approaches the problems of maintaining effective training programs. Follow-up conversations were held with appropriate personnel to verify interpretation of the documentation and to obtain information regarding evaluation practices. Visits were made to Headquarters, Air Training Command; the United States Air Force School of Applied Aerospace Sciences (Lowry Air Force Base, CO); Naval Education and Training Command; Naval Technical Training Command Headquarters; and the United States Army Training and Doctrine Command. Training Appraisal Plans prepared by Navy training activities were obtained from CNET (N-34) and reviewed to determine personnel and organizational capabilities for the conduct of training evaluation. The review was conducted between June 1975 and May 1976.

Section II of this report describes the various military training evaluation efforts. Section III briefly summarizes significant features of the programs and offers evaluative comments.

## SECTION II

### CURRENT MILITARY TRAINING EVALUATION PROGRAMS

Each military service prescribes evaluation requirements and programs for assessing the quality of its training. There is considerable variability in underlying philosophies and in specific mechanisms employed for training evaluation. This section describes training evaluation programs currently conducted within the military.

#### U.S. AIR FORCE TRAINING EVALUATION

A formal training evaluation program has been continuously in effect within the Air Force since 1960. This program provides a high degree of control over the training system. Evaluations of training are conducted on a formal basis and certain other mechanisms are available to the Air Training Command (ATC) for assessing operation of the training system. Training evaluation is monitored at Headquarters ATC. Objective Status Reports are prepared there which summarize any deficiencies observed in any training course for which ATC is responsible. Deficiencies of courses are reported (at least) monthly to the Commanding General of ATC along with identification of corrective action to be taken, in process, or completed.

PURPOSE OF PROGRAM. Air Training Command Regulation 52-1, Training Evaluation and Course Reviews (1974), establishes training evaluation requirements. The Air Force Technical Training Evaluation Program includes the "collection, collation, analysis, and interpretation of feedback information to assess the effectiveness of training courses and the extent to which course graduates satisfy field performance requirements." The goal of this program is to provide both internal and external feedback for the improvement of Air Force instructional systems. Evaluation is made of all military, technical, and career development courses.

Information which is obtained from internal sources includes student attrition rates, student acceleration, instructor observations and evaluations, student critiques, student counseling, and reviews of training materials. External sources of feedback information include field evaluation visits, graduate and supervisory questionnaires, job performance evaluation, coordination of training standards, command comments, Training Quality Reports, Career Development Courses (CDC), failure rates, student comments on CDCs, occupational surveys, and training advisors.

EVALUATION RESPONSIBILITIES. Five levels within the Air Force organization are involved in the formal training evaluation program:

TAEG Report No. 39

- . Air Training Command Training Evaluation Division
- . Training Schools
- . Training Evaluation Divisions at the schools
- . School Operations Divisions
- . Training Groups

The Air Training Command Training Evaluation Division (ATC/TT) establishes technical training evaluation policies, guidance, and programs. Surveillance over the technical schools' evaluation programs is maintained by ATC/TT which reviews evaluation plans and reports submitted by the schools. This group also prepares briefings and reports for the ATC Commanding General regarding training course status. The ATC/TT also coordinates with major commands on field evaluation problem areas and maintains liaison with other services and the civilian academic sector regarding new evaluation concepts. The ATC/TT conducts seminars and workshops for the improvement of evaluation programs and directs special evaluations as required.

Training schools implement training evaluation and course review programs and approve training evaluation plans and reports, and field evaluation visits. Training Evaluation Divisions (TTE) at the various Air Force schools establish school evaluation priorities, conduct evaluation in accordance with approved training plans, conduct field evaluation visits, prepare evaluation reports, and maintain staff surveillance over the course review program. The TTEs evaluate courses which award Air Force Specialty Codes (AFSC) at least once every 3 years. Special evaluation may be conducted on any course at any time if more immediate feedback is required (e.g., course revisions, outside reports of deficiencies). The TTEs are typically staffed by a military director, several civilian and military training specialists, and a number of clerical and data processing staff. Currently, the TTEs at the schools report to the school Commander in parallel with the training departments. It is understood, however, that consideration is being given (at ATC Headquarters) to amending this procedure. The proposal is to have the evaluation group report directly to the Training Center Commanding General.

School Operations Divisions (TTO) assist in the identification of evaluation requirements and conduct/participate in internal evaluations. These divisions also coordinate feedback and planning actions with the Training Evaluation Division when new courses are established or major course revisions are made. The Training Departments at the schools schedule and conduct resident course reviews. They also participate in

planning and conducting evaluations and take corrective action as indicated in Training Evaluation Reports.

The Air Force Training Evaluation Program contains provision for checking graduates' performance after they have left the school (external evaluation) and also for checking the quality of instructional elements within the school (internal evaluation). External evaluation is accomplished principally through the Graduate Evaluation Program, but there are also additional checks on training quality. Internal evaluation is accomplished through annual course reviews and student critiques. These programs are described next.

GRADUATE FIELD EVALUATION PROGRAM. Graduate field evaluation is a primary tool in the quality control of formal and career development courses. This program provides information concerning:

- . The ability of graduates to do their job to assigned proficiency
- . The extent to which acquired skills are used by recent graduates
- . The extent to which knowledge attained is retained by recent graduates
- . The need to revise Specialty Training Standards (STS) or training courses
- . The need for further evaluation of training problem areas identified by the evaluation of graduates.

Field evaluation of formal school graduates involves: field evaluation visits, graduate and supervisor direct correspondence questionnaires, job performance evaluations, and training quality reports. Field evaluation visits and graduate and supervisory questionnaires are the major information sources for the graduate evaluation program.

Determination as to whether an individual can perform his job "to assigned proficiency" is made by reference to the STS for a particular Air Force Specialty. These STSs, prepared for each AFSC, become a permanent part of an Airman's personnel file. The STS lists the tasks that the holder of the AFSC should be able to perform and also identifies the level of skill and knowledge he should possess for the job level which he holds. Thus, an objective checklist is provided evaluators for use in conducting individual evaluations.

The Graduate Evaluation Program under which course effectiveness is determined requires the completion of a number of sequential steps. Initially, a plan is developed which outlines the objectives of the evaluation, the methods that will be used, and the schedule for task completion. Specific responsibilities for evaluation are also assigned.



Questionnaires for obtaining responses from both course graduates and their supervisors are prepared and mailed. Typically, questionnaires are mailed within 3 to 6 months after a student has graduated from the course. Returned questionnaires are processed and analyzed to determine training effectiveness. If 80 percent (or more) of the course graduates surveyed accomplish a training objective at or above the prescribed training standard skill level, the training course is considered to be satisfactory. Results of the data analysis are subsequently incorporated within a Training Evaluation Report (TER) which also outlines recommended actions to remedy any observed training deficiencies. The Training Evaluation Division at Headquarters ATC reviews all TERs and prepares summaries relating to individual courses. Courses which have exhibited training effectiveness deficiencies are reported to the Commanding General at monthly briefings. These become items in the ATC Management-by-Objectives (MBO) program. Recommended actions to alleviate deficiencies, status, and "get-well" dates are included. The Training Evaluation Division and the Training Department coordinate follow-up action to insure that the training evaluation report recommendations are in fact implemented. Follow-up evaluation to determine whether the implemented recommendations actually improved the effectiveness of the particular course is generally not conducted.

The Graduate Evaluation Program also has provision for field evaluation visits which may be made to selected bases within 6 months after graduates are assigned. These visits are generally made whenever there is a need for additional information from course graduates, their supervisors, or field managers. The purpose of the field visits is to obtain information regarding the frequency of use of skills and the ability of the graduates to perform the tasks for which they were trained. The number of field visits actually made, however, is small because of funding limitations.

**OTHER CHECKS ON GRADUATE QUALITY.** In addition to the formal evaluation procedure described above, the ATC system contains provision for other checks on graduate quality. These are: (1) Training Quality Reports, (2) Inspector General (IG) Inspections, and (3) Job Performance Evaluations.

Training Quality Reports. The Training Quality Report (AF Form 1284) provides a quick reaction evaluation capability for determining if the training system is performing its functions. Its use, required by Air Force regulation, enables command personnel to report on an immediate basis if: (1) the graduate(s) does not meet the proficiency level specified for a task or knowledge as listed in the approved STS, (2) the graduate(s) is not required to perform tasks listed in the STS while working in his assigned AFSC, or (3) the STS code levels or tasks exceed the requirement of the graduate's AFSC. Thus, overtraining or under-training, as identified by supervisory personnel, can be reported directly to the ATC/TT. These reports are also sent to the Evaluation Center at the affected school, Air Force Headquarters, and the affected Air Command. Corrective action may be taken on the basis of these reports. In the past,

submission of unfavorable Training Quality Reports has also triggered ATC IG investigations to determine the locus of deficiencies at training centers.

Inspector General Inspections. The ATC IG team provides an external check on the ATC training system. Scheduled for 18-month intervals, the ATC IG inspection (which is a separate function from the AF IG inspection) examines all facets of operation of a school. Specialists in various fields (e.g., financial specialists, training specialists) on (typically) 2-year assignments compare practice at the schools with established ATC standards. Deficiencies are noted and corrective action recommendations identified. Reports sent directly to the ATC Commanding General also cover the issue of training effectiveness.

Job Performance Evaluations. Job performance evaluations can provide still an additional check on training quality. When performed, they are conducted by individuals who are knowledgeable of the particular job assignment. This involves observation of a graduate's performance to determine if he can do the job adequately. This type of evaluation is costly and time consuming and is typically conducted only if the required training evaluation information cannot be obtained in any other way.

COURSE REVIEW PROGRAM. The Air Force Course Review Program is the formal review process through which all technical elements of an instructional system are examined annually to determine if they are current and effectively supporting the training objectives. Self-examination checklists are used by the schools to assess adequacy of, for example, course control documents, training literature, training equipment, training aids, facilities, instructional methods and techniques, student measurement, faculty and supervisory staff, and student critiques. Recommendations for improvements to the course are made from the course review results and follow-up action is taken. The annual course review is conducted by the school's training department.

Standardization/Evaluation (STAN/EVAL) teams also provide a check on the quality of training courses conducted by ATC. At present, STAN/EVAL teams are located at each Air Force training center. These teams work under the guidance of the ATC STAN/EVAL section and report to the Vice Commander of the training center at which they are stationed. Unlike the ATC IG team which examines all facets of operation of a school, the STAN/EVAL team focuses assessment attention on training only. Internal evaluations of training are conducted by this team on a recurring basis. Selected courses (usually high flow courses) are evaluated annually. Six functional areas are examined by the military training specialists who comprise the team:



## TAEG Report No. 39

- . Management and supervision
- . Curricula
- . Instruction
- . Training aids and equipment
- . Facilities
- . Instructor program

Reports made to the Vice Commander at the training center specifically note those areas where ATC standards are not being met. STAN/EVAL reports are also sent to Headquarters ATC. Programs for correction of observed deficiencies become part of the Commanding General's MBO program.

STUDENT CRITIQUE PROGRAM. The ATC student critique program also provides evaluative information about training courses. This program is designed to obtain constructive criticism from students regarding academic instruction received in training courses. Critiques may be submitted by individual students or by an entire class of students. Critiques are routed to the appropriate department (to the training department in the case of a critique of a course) for action consideration. After review, feedback regarding disposition of the critique is given to the individual(s) who submitted it. A random sample of student critiques is reviewed by the Training Evaluation Division and/or the Center or School Commander on a periodic basis. If relevant, information from student critiques is also included in TERS.

### U.S. NAVY TRAINING EVALUATION

Several significant training evaluation programs are in effect within the Naval Education and Training Command (NAVEDTRACOM). In general, the flavor of Navy policy is to provide support for evaluation to those who conduct training. Policy places much less emphasis on control of the training system than it does in the Air Force. General guidance regarding requirements for evaluation programs is promulgated by the CNET and his Functional Commanders. Individual schools, conforming to the general Command guidelines, develop their own specific programs to accomplish evaluation of their training courses. Evaluation programs which have application NAVEDTRACOM-wide are described below. Certain other programs which impact on Navy training evaluation are also reviewed.

At present, three documents pertain most directly to the evaluation of Navy training courses--two CNET Instructions and the Chief of Naval Technical Training (CNTECHTRA) AIO Manual. CNET Instruction 1540.3 establishes the necessity for and the basic requirements for conducting

evaluations to improve the effectiveness of technical training. CNET Instruction 1540.6 builds upon these basic requirements and provides instructions for the establishment of an organization within training units (i.e., a Curriculum Instructional Standards Office (CISO)) to conduct training effectiveness evaluations and to assist in the development of training content, methods, and media. The CNTECHTRA A10 Manual provides procedures and instruments for use in conducting evaluations of training. The recently promulgated instructions for course design using Instructional Systems Development (ISD) procedures (NAVEDTRA 106A) also contain provisions for evaluation of courses designed in accordance with ISD procedures.

CNET INSTRUCTION 1540.3, APPRAISAL AND IMPROVEMENT OF TRAINING. The stated purpose of CNET Instruction 1540.3 is to provide information for systematizing training appraisal for improvement of training and training effectiveness. The goal of the training appraisal program is "to provide an objective determination of the quality of the output of the training system and to provide the means for correcting the quality when there are deviations from prescribed standards."

The instruction requires both internal and external appraisals of training courses. Internal appraisal is based upon information obtained within the school or course. The following are specified to be prerequisite to conducting an effective internal appraisal:

1. Learning objectives based on task analysis
2. Criterion measures of student performance
3. Effective analysis/use of student test data
4. Effective procedures for corrective action/follow-up evaluation of changes
5. Supervisory support/effective administrative procedures and regulations.

The instruction suggests that the development of a single internal evaluation plan for use by all CNET activities is not practical at this time due to the diversity of training within the NAVEDTRACOM. Therefore, each school is directed to develop its own training appraisal plan to include at least the following elements of internal evaluation:

1. Review of course documents to determine if discrepancies exist between the planned instructional system and what actually occurs

TAEG Report No. 39

2. Review of resources (such as training facilities, equipment) to determine if they are adequate
3. Observation of classroom activities (to include review of training aids and materials)
4. Evaluation of instructors' performance
5. Review of student measurement program
6. Monitoring trends in training statistics such as attrition, setback, etc.
7. Well designed student critique program.

External evaluation is not addressed in detail. It is described as being conducted to determine both how well course graduates can perform the job and the degree to which course learning objectives are relevant to the job requirements. This information is to be obtained through the use of graduate questionnaires, supervisor questionnaires, ship visits, task analysis, and letters from Fleet Commanding Officers.

CNET INSTRUCTION 1540.3A (PROPOSED). CNET Instruction 1540.3A, Appraisal and Improvement of Training, currently in preparation, will cancel CNET Instruction 1540.3. The purpose of 1540.3A is "to provide information for standardizing training appraisal and to provide guidance for establishing training appraisal plans within the Naval Education and Training Command (NAVEDTRACOM)."

Essentially, the provisions of 1540.3 are retained. The most significant change is the formalization of procedures for obtaining feedback data from the Fleet for use in training effectiveness determination. Largely, the use of questionnaires is advocated for this purpose. The questionnaire method to be used is based on the results of a TAEG study (Dyer, Ryan & Mew, 1975) which compared several means of obtaining post-formal training feedback information. It was found that a well designed questionnaire provided data (Radioman "A" School) that could be used to identify training problems at the school. Information obtained by questionnaire was nearly identical to that obtained via face-to-face interviews. A second report (Dyer, Mew & Ryan, 1975) provides detailed instructions for preparing and using this type of questionnaire.

CNET INSTRUCTION 1540.6, TRAINING APPRAISAL SUBSYSTEM. CNET Instruction 1540.6 provides guidance regarding the establishment at training activities of Curriculum and Instructional Standards (CIS) Offices or Departments to:

- . Provide advice to Commanding Officers regarding the efficiency and effectiveness of prescribed education and training programs
- . Maintain prescribed standards for curriculum design, testing, instructor performance, and training aids
- . "Conduct formative evaluation of training: (1) internal evaluation to determine training efficiency and (2) external evaluation to determine training effectiveness including training transfer."

The CIS Offices are to maintain quality assurance of training within prescribed standards through review of curricula, documentation, classroom monitoring, guidance, inspection, and maintenance of publications and training aids, and instructor staff inservice training programs. The methods to be employed in such reviews are not specified. Similarly, the testing program is to be maintained "within prescribed standards" through construction and validation of tests, analysis of test data, etc. Also, feedback data is to be collected, processed, evaluated, and reported on, regarding training quality using methods of questionnaires, student critiques, structured interviews, etc. More precise exposition of these methods is not offered. The CIS Offices/Departments are to be headed by a Special Assistant who reports to the Commanding Officer of the training activity. The CIS departments are to be administratively separate from the instructional departments but provide them technical assistance/expertise in the accomplishment of quality training.

The typical CIS organization consists of two branches--an evaluation branch and a curriculum and training support branch. The evaluation branch is charged with responsibilities such as the maintenance of item banks for preparation of examinations for use by the instructional departments, statistical analyses of test data for use in training evaluation and test item improvement, internal and external feedback, development of proposals for short- and long-term evaluation projects, preparation of Training Appraisal Plans, and student critique programs. The other branch is involved in a wide variety of activities including task analysis, curriculum development, monitoring classroom instruction, instructor inservice training, monitoring procurement of training devices, development of training aids, and maintenance of a central technical library.

CNTECHTRA A10 MANUAL. The Procedures for the Planning, Design, Development and Management of Navy Technical Training Courses, the CNTECHTRA A10 Manual, was developed under the direction of the Chief of Naval Technical Training. It contains procedures to meet unique needs in planning, design, development, and management of CNTECHTRA courses. NAVEDTRA 106A, Interservice Procedures for Instructional Systems Development, (ISD), is acknowledged as the basic publication for ISD. The A10 Manual,



however, is designed for specific requirements of the Naval Technical Training Command. A revision of the A10 Manual to incorporate the model, terminology, and other requirements of NAVEDTRA 106A is in process. The section of the Manual which deals with "Management of Navy Technical Training Courses" addresses both internal and external evaluation.

Internal Evaluation. According to the A10 Manual, the purpose of an internal evaluation is "to determine that the elements of the course are current and are working effectively and efficiently to achieve the Learning Objectives." Internal evaluation is further described as assuring that: (1) the learning objectives are based on task analysis, (2) accurate and appropriate criterion measures are provided, (3) effective use is made of student data, and (4) efficient and effective supervisory support is provided. Internal course evaluations, referred to as "Course Reviews," are to be conducted annually.

Procedures. The A10 Course Review procedure consists of staff personnel completing checklists in each of the following areas:

1. Course control documents
2. Testing
3. Instructional Staff
4. Instructional Materials
5. Course Plans and Data

A number of review items in each of these areas is rated according to the following scale:

Adequate - necessary requirements being met

Generally Adequate - minor attention/improvement needed

Inadequate - major attention/improvement required

Not Applicable - review item does not apply

At the end of each of the five sections, space is allotted to explain each of the "Inadequate" ratings and to make recommendations for corrective action. The manual specifies that the individuals who complete these evaluation checklists should be "trained" personnel who not only have knowledge of the course but also have (1) daily contact with the course, (2) wide teaching experience in the course as well as recent Fleet experience, and (3) if possible, assisted in the design/development of the course.

Utilization of Data. The completed course review checklists are reviewed by "qualified personnel" who prepare a summary report to the Commanding Officer of the training activity. This report summarizes the evaluation data upon which recommendations for improvements in training are based. No procedure is specified by which the recommendations for improvements are followed up and/or evaluated to determine if the deficiency was corrected.

External Evaluation. The A10 Manual describes two types of external evaluation data: summative and formative. Summative data is defined as being information about how well course graduates can perform on-the-job. Formative data is distinguished as being information about how relevant the course learning objectives are to the knowledge and skill requirements of the job. (These terms have somewhat different meanings in other contexts.)

Procedures. Five methods of obtaining external evaluation data are described:

1. Questionnaires administered to graduates on the job
2. Questionnaires administered to supervisors of course graduates on the job
3. Visits to the job to observe graduates' job performance
4. Analysis teams which make surveys and perform analyses regarding job requirements in particular ratings or occupational fields
5. Unsolicited feedback from staff in the field regarding the adequacy of training.

ISD EVALUATION PROGRAM/CONCEPTS. The recently promulgated ISD procedures (NAVEDTRA 106A) which will be applied to the design of future Navy training courses also require that evaluations be conducted. Internal evaluation as specified by the ISD model consists essentially of checking to insure that the established procedures are adhered to and applied in the intended ways. It is tacitly assumed that the procedural steps are correct and that if correctly applied the resulting course will be satisfactory.

Thus, the primary purpose of the ISD internal evaluation program is to determine if a course has been developed/conducted according to the "standards" specified in the ISD procedures. The ISD internal evaluation model is intended to determine if a course of instruction provides students with learning experiences appropriate to achieving the course learning objectives. In theory, the characteristics and needs of the students are assessed to determine the type and form of instruction



appropriate to them so that courses can be tailored for particular populations. For internal evaluation, data are collected regarding student achievement and used to revise aspects of the course which exhibit deficiencies. The primary goal of the ISD external evaluation program is to determine whether the course learning objectives are relevant to the actual job requirements.

Internal Evaluation. In the ISD context, internal evaluation includes the determination of whether the instructional development effort has accomplished what was intended. The following procedures are featured:

1. Development of a progress evaluation plan to enable managers to assess the adequacy of the progress of the ISD effort
2. Development of a process evaluation plan to describe and document the actual developmental process for the particular course being developed
3. Development of a performance evaluation plan to determine students' external requirements (i.e., their qualifications for the instruction), entry skills, performance on internal tests and time required to complete instructional units
4. Development of a plan for collecting information from students about the perceived quality and preferability of instructional events and materials
5. Development of a plan for collecting information from instructors regarding, for example, problems students have with particular course objectives, time spent in presenting instruction, and opinions about instructional materials and procedures.

External Evaluation. To conduct an external evaluation, according to the ISD Model, a plan is developed which specifies what actions will be taken to obtain the required information about the graduate's ability to perform on the job. Data may be obtained from course graduates and/or their supervisors. Information that can be obtained from course graduates includes their opinions about:

- . How well they believe they can perform the job
- . The kind and amount of training received since arriving on the job
- . How well the instruction prepared them for the job

TAEG Report No. 39

- . The portions of the course which were relevant to the job
- . The tasks which cause them the most difficulty.

Information that can be obtained from the supervisors of course graduates includes:

- . How well the graduates are performing the job
- . How those graduates compare to those who received another form of training
- . Areas in which the graduates were not adequately prepared.

The manual stipulates that other information can be obtained from other sources. Ideally, evaluation teams can provide data concerning how well course graduates score on job performance measures obtained either through observation or by special testing. They can also provide information concerning which job performance measures were the most difficult. Information about task performance is to be summarized and analyzed and recommendations made for revisions or changes to the course and/or the job analysis process.

PERSONNEL TRAINING AND EVALUATION PROGRAM (PTEP). One of the more significant evaluation programs within the Navy is that conducted by the submarine community within the Fleet Ballistic Missile (FBM) Personnel Training and Evaluation Program (PTEP). At present, this program is modest in scope and is principally confined to the assessment of the job knowledge of a small, but growing, number of maintenance ratings. Plans are to expand the program outside the FBM community.

The PTEP was established initially to provide a mechanism for evaluation of the FBM Weapons System Training Program. Two components of the FBM training system are important to the achievement of PTEP goals: Personnel Performance Profiles (PPP) and the Training Path System (TPS). The PPPs, typically developed by contractors, consist of descriptions of the knowledges and skills needed to operate and maintain specific equipments or systems. The TPS specifies the profile items and levels of achievement appropriate to (i.e., the training requirements for) each FBM Navy Enlisted Classification. Thus, they provide an objective listing of what an individual should know and be able to do to perform effectively. In providing a listing of job requirements, they serve a function similar to the Air Force's STSs.

Test scores provide a means for selection and assignment of personnel, for determining training needs, and for evaluation of personnel capabilities. Each individual within an affected rating is tested twice annually to determine his knowledge of operations required to maintain equipments

for which he is responsible. Tests are changed every 5 months. Test items are drawn from a pool of standardized items. Those items in turn are based on an analysis of the maintenance needs and characteristics of the equipment which defines skills and knowledges required to maintain the equipment. Strictly speaking, the evaluation portion of the program is not concerned directly with training evaluation but is, instead, concerned with personnel evaluation. However, the test program is adaptable for use in training evaluation.

Two basic types of tests, differing principally in their degree of comprehensiveness, are prepared and administered by the PTEP organization. Course Achievement Tests (CAT) are administered at the completion of a particular course of instruction. System Achievement Tests (SAT) are administered to individuals who have been on the job for some period of time. The CAT provides a suitable device for assessing training effectiveness especially if used in conjunction with a standardized pretest. Given at the end of a course of instruction, scores on this standardized test could be used to determine how much learning (of job knowledge) the course has brought about; i.e., to assess training effectiveness. SATs are designed to measure the proficiency of individuals relative to the knowledge and skill requirements specified in the PPP and TPS for particular items of equipment.

Test scores may be used for personnel assignments, for example, to insure that at least one member of a crew will possess adequate knowledge of some particular piece of equipment. In addition to this "composition of crew" use of test scores, they may also be used to identify individuals who require training. They have also been used, in at least one instance, to identify unnecessary training. Here, scores of individuals having undergone a particular advanced course of instruction were compared to those not having received the formal instruction. No differences were found in test scores nor were there any differences in equipment casualties aboard their respective ships. Hence, the conclusion that this particular advanced course could be eliminated which would result in a training cost saving.

At present, no formal mechanism exists for using PTEP scores for improving training courses. Affected schools are required to administer the CATs so that the scores will be available but are not required to use the examination results. However, informal reports of PTEP staff to instructors and other training personnel detailing areas of noted student underachievement frequently result in alterations to courses.

#### U.S. MARINE CORPS TRAINING EVALUATION PROGRAMS

No formal, systematic program for determining the effectiveness of Marine Corps training was uncovered during the course of this review. The evaluation provisions that exist are incorporated within documents

written for other purposes; e.g., to regulate training within a larger scope. Salient evaluation provisions of selected Marine Corps documents are presented below.

ACADEMIC REGULATIONS (ECO P5000.IJ). The stated purpose of the Academic Regulations is to implement the Marine Corps' system for design of courses of instruction and to prescribe academic regulations within the Education Center for the conduct and support of instruction. The regulations stipulate that "School Directors/Commanding Officers are responsible to the Director, Education Center, for the total performance of a school in accomplishing its educational mission." In the area of evaluation, specific responsibilities are that they implement and supervise a comprehensive evaluation program. The program must include:

- . An active testing program, for which the goal must be to test each learning objective, utilizing the criterion-referenced test method as primary evaluation vehicle.
- . The use of both formative and summative testing, to the greatest extent possible, in order to collect a variety of useful information for educational decision-making.
- . Provision for the use of Instructional Rating Forms (IRF) which are to be completed for each lesson presented.
- . Provision for the use of After Instruction Reports (AIR) which are to be submitted for each lesson presented.
- . Provision for using Graduate Questionnaires to obtain feedback from former students and their supervisors.

Evaluation and Validation. Evaluation and validation is identified as a major component of the Marine Corps system for design of courses of instruction. Evaluation of a given segment of instruction includes (1) student responses on written critiques and in-conference group discussions, (2) questionnaires submitted by graduates and their supervisors, (3) analysis of test results, and (4) supervisory evaluation of the conduct of instruction.

Internal evaluation procedures specified by the Academic Regulations include annual reviews of all lesson plans and periodic reviews of job/task inventories and instructional strategies/concepts. The graduate questionnaire program is the primary means used for validating the instructional system. This program is designed to obtain information about actual field performance of the graduates, and this feedback is to be used for decisions concerning the appropriateness of instructional objectives. Observed deficiencies serve as a basis for revising course content or instructional methods.



Educational evaluation is defined as involving the application of judgments and standards to make decisions for improving the educational system. Evaluation is based on the analysis of data collected in response to specified information requirements. Decisions are to be based on information that compares performance with goals and preselected standards. Accordingly, an effective evaluation program is viewed as one which includes:

1. Identification of the decision(s) to be made
2. Determination of the type of information needed to make the decision(s)
3. Collection of appropriate information
4. Assessment of information against appropriate standards
5. Application of value judgments to make the decision(s).

Evaluation Documents. Tests, Instructional Rating Forms, After-Instructional Reports, and Graduate Questionnaires are described as means for obtaining evaluative data.

Tests. Active testing programs are specified to promote and assess student learning and instructional effectiveness. Courses of instruction are considered effective if 90 percent of the students master 90 percent of the course objectives (90/90). Instruction is considered unacceptable if the results fall below 80/80. Students may be given pretests, progress tests, posttests, and retention tests. Both formative and summative testing are to be used in evaluating instruction.

Criterion-referenced tests are to be used to test each learning objective. The Academic Regulations stipulate that student "competence will not be judged by comparative achievement levels. There is no room for a relativistic approach in a professional Marine Corps." Since the primary thrust of the testing program is for purposes other than grading, sampling techniques may be employed for determining instructional effectiveness. A representative, random sample of students (usually not less than 10 percent) may be given formative tests to provide information about the effectiveness of the instruction. Student input to the evaluation process is also provided more directly via Instructional Rating Forms (IRF).

Instructional Rating Forms. At the Marine Corps Education Center, IRFs are completed by students after each lesson is presented. In theory, these forms provide data "that allows decisions to be made about the curriculum, the instructional process, support activities, lesson related materials, and tests." Students record their impressions about the relevancy of lesson objectives to their future job, difficulty of the material, utilization of class time, and value of training aids. Ten percent of the class, or five students (whichever is greater), are required to complete the forms.

After Instruction Reports. In addition to being rated by students, the instructional process is also "evaluated" by instructors. After each lesson is presented, an After Instruction Report (AIR) is completed by the course instructor. The instructor may review the Instructional Rating Forms (student ratings and comments) for incorporation of significant items within the AIR. The instructor is to assess all factors which affect the period of instruction. When identifying deficiencies, he is to indicate the type and level of intervention he feels necessary to upgrade the lesson in its next cycle. The AIRs are reviewed and commented on at higher levels of the training organization and become a primary management tool for controlling training quality.

Graduate Questionnaires. The use of questionnaires to solicit information about graduates' job performance is required by the Academic Regulations. Former students are requested to provide feedback concerning the value of the educational experiences they gained while in the school after they have had an opportunity to function in a field situation. Similarly, supervisors of graduates are also solicited for feedback concerning the degree to which graduates have retained and used the objectives featured in the school program. The obtained feedback information is used to revise course learning objectives and/or for general improvement of training (e.g., updating task analysis data).

DESIGN OF COURSES OF INSTRUCTION (MCO P1510.23A). The stated purpose of this document, Design of Courses of Instruction, is to publish guidance for the development of formal courses of instruction. The manual specifies procedures and administrative requirements for instructional design using a systems approach. It stresses the importance of obtaining and using job task information for the design of training courses. Job task information is to be used for the development of training objectives upon which courses will be based. In turn, achievement of the objectives is to be used as the basis for determining instructional effectiveness.

Evaluation procedures; i.e., testing, is required by the manual. The stated purpose of testing is:

1. To pretest a group of students
2. To evaluate, graduate, or eliminate students from the instructional program
3. To diagnose learning difficulties
4. To maintain quality control, and
5. To measure the adequacy of the instructional system, identifying the weaknesses, and forming a basis for modification.



Four types of tests--performance, written, oral, and ratings--are identified. Performance tests are considered to be the most desirable since they require the student to demonstrate a learned behavior. Criterion measurements are to be used to measure the learning objectives prescribed for the course of instruction. The emphasis of testing is placed on objectives which have been identified as the most important for job performance. The manual stipulates that criterion measurement, once developed, will be administered to a pilot group to determine if they are valid, reliable, objective, economical, administrable, and standard.

Internal Evaluation. The manual presents a number of "normally accepted means of evaluation" to serve as guides for commanders conducting internal evaluations of formal courses of instruction. These include the use of school evaluators, instructor evaluation, student evaluation, instructor and student interviews. Faculty or staff personnel familiar with the objectives of evaluation, the objectives of the training, lesson plan content, and the component requirements for the course of instruction are identified as the proper personnel to evaluate classes.

Internal evaluation practices include the collection of comments about and ratings of course material from both instructors and students. At the end of each block of instruction, instructors are to record discrepancies noted by them during the conduct of instruction. Written student comments concerning the learning objectives, course content, strategies, and testing are also obtained. Private interviews may be held with students and/or instructors to supplement written comments.

Validation. Validation refers to what the Navy describes as external evaluation and the Air Force calls graduate evaluation. The principal purpose of validation is to assure that the course of instruction is effective and that it produces the desired results. The need for revision (i.e., additions, deletions, or improvements) of courses is noted whenever there are discrepancies between planned and actual training outcomes.

Feedback Evaluation. One of the principal means of course validation is the feedback information that is received via questionnaires from graduates and their supervisors. Commanders conducting formal courses of instruction are required to accomplish feedback evaluation procedures. Guidelines are provided for the programs. They prescribe that:

1. Each course will have a questionnaire, or questionnaires, designed to provide feedback information to the school director. Information solicited is to consider both the effectiveness of the instruction presented and also its appropriateness. The concern is to determine whether (a) the course learning objectives were achieved, and (b) whether the course learning objectives support the requirements of the field commander.

2. The principal source of feedback will be from supervisors of course graduates on the next assignment following completion of the course. Questionnaire design should facilitate "critical, objective evaluation of the graduate's ability to apply the learning gained in the course." Points for evaluation should be precisely stated and specific duty task skills as related to school learning should be identified.

3. Questionnaires can also be used to obtain data from graduates. Questionnaires to be answered by the graduate "should encourage constructive, critical appraisal of the adequacy, practicability, pertinence, etc., of such aspects as facilities, course content, and instruction."

The following methods, or combinations thereof, are also considered valid and appropriate (by the Design of Courses of Instruction) for obtaining feedback data on course effectiveness:

"a. Inclusion of information by endorsement and enclosure on the orders of graduates when detached;

b. Insertion of a request for the information in the service record or qualification record of graduates when detached;

c. Blanket letter requests, with pertinent information, to commanders of organizations to which graduates have been assigned."

MARINE CORPS COMMAND AND STAFF COLLEGE EVALUATION PROGRAM. The Marine Corps Command and Staff College (C&SC) also generally applies the evaluation procedures identified in the previously described governing documents. IRFs and AIRs are used to obtain information for the quality control of the process of instruction. Interviews may also be conducted with students to obtain information for formative evaluation of instruction. The input gathered from interviews is inserted into the validation and evaluation process by means of comments in the AIR. Questionnaires are mailed to graduates and their supervisors 6 to 8 months after course completion to obtain information about field performance. The returned questionnaires are reviewed annually to glean inputs for training course revision. In addition, an active testing program is described. Test results are to be used to assist in directing the student's efforts and to make appropriate revisions to lesson materials.

MARINE CORPS UNIT LEVEL TRAINING MANAGEMENT (MCO P1510.26). The Unit Level Training Management Order assigns responsibilities to unit commanders to apply the systems approach to devise training for "raising unit or individual performance from current levels of performance capability to desired/required levels." Evaluation tasks are included. The style of the manual (MCO P1510.26) is largely tutorial. It provides an explanation of terms, purposes of evaluation (including improvement of the content and methods of instruction), and types of "tests" (i.e., written/verbal, performance and observation/evaluation) that can be used for evaluative purposes. Performance tests are identified as the "most appropriate for use at the unit level," and guidance is provided for the preparation and

## TAEG Report No. 39

conduct of such tests. The guidance includes a description of the characteristics of good tests, the need for test validation, and how to administer such tests.

MARINE CORPS' GENERAL TRAINING SYSTEM (GENTRAS). Detailed information regarding the current status of the Marine Corps' General Training System (GENTRAS) was not available for review during this project. A description of the system as it was conceived in 1970, however, included several significant features for the quality control of training. Included in GENTRAS is a testing and student evaluation subsystem which, in concept, can be used to determine apparent deficiencies in training. From this, appropriate recommendations for training improvements can be made to the responsible schools.

A "Ratings File" was envisioned for inclusion within GENTRAS. This file would contain data on individual students for classes surveyed in a field evaluation process. Classes would be surveyed on a sample basis. Information to be entered into the file would include ratings by supervisors of the performance of recent graduates (i.e., within 6 months after graduation from a surveyed course). Ratings would be collected by questionnaire, and individuals would be rated on each skill taught in the course. The information could then be compiled to arrive at a Field Proficiency Rating (FPR) which would show the composite performance of all students rated on each skill. A low FPR for a skill would indicate that it is probably not taught effectively. The average proficiency of each student performing all skills could also be determined as could a student's field rank. The amount of agreement between the student's field rank and his rank in class was seen as useful to "provide some insight into what effect performance in the course had on performance in the field."

### U.S. ARMY TRAINING EVALUATION

A number of major Army schools conduct individual training in accordance with general guidance and policy established by the Training and Doctrine Command (TRADOC). The schools establish specific programs for conducting training and for evaluating its quality within these established guidelines. Due to constraints on the study, it was not feasible to visit individual schools to ascertain the nature and quality of the training evaluation programs in effect at them. Consequently, this review is limited to the general Army-wide concepts, practices and policies affecting training evaluation and the determination of training effectiveness.

The Army training system is currently undergoing a number of changes. The most significant is that the role of the formal school as the principal personnel training agency is being gradually shifted to the individual's unit of assignment. It is anticipated, probably within the next 5 years, that institutional training (i.e., schools) will provide only core type training. Individuals will then receive the larger part of their necessary job training within the unit to which they are assigned.



The schools will, however, continue to participate in the development of training and evaluation materials for unit use.

In the present Army training context, evaluation of both individual and unit proficiency is accomplished principally for diagnostic purposes; i.e., to determine where weaknesses lie so that they may be corrected by additional training. It is understood that at the present time, individual schools do relatively little to assess the effectiveness of the courses they conduct. Formal external evaluation programs for use in training course improvement are not a standard feature of the Army training system.

ARMY TRAINING AND EVALUATION PROGRAM (ARTEP). Within the current Army concept, the effectiveness of training is judged in terms of the contribution of training to a unit's mission readiness/accomplishment. Readiness is determined by testing in a staged battlefield exercise. Performance objectives are derived by analysis of the requirements of the specific missions that the units might be required to perform. The specific performances, the conditions under which they will be performed, and standards are identified. The ability of the unit (and individuals within the unit) to exhibit the required behaviors is observed and checked by evaluators to provide an objective determination of skills. Score cards are prepared for squads, platoons, companies, etc. No aggregate scores are given since Army interest is not in evaluation per se. Officially, ARTEP is considered to be a diagnostic tool for Commanders, and evaluation reports go no higher than the Division Commander.

When the unit cannot perform to required standards, this is considered to be a training deficiency; i.e., training has not been effective. No deliberate attempt is made to attribute the training deficiency to any particular prior training experience (e.g., a school course). Training is viewed in a total context with all training combining to produce the desired performances. If the performances are not as desired, then additional training is given.

The development of ARTEP accompanied a revision in the Army's training management doctrine. In November 1971, the Army abolished the requirement for mandatory training and decentralized responsibility and authority for the management and conduct of training (TMD-1). This decentralization shifted the focus of the Army training effort to the unit level where the job is actually performed. With this shift an increased emphasis was placed on performance-oriented training. This resulted in a movement away from evaluating students' acquisition of subject matter to evaluating the achievement of performance-oriented training objectives. Under the decentralized training philosophy, the authority and responsibility for the planning, conduct, and internal evaluation of training have been delegated to battalion and separate company commanders. "...the determination of specific training objectives has been left largely to the commander most familiar with his soldiers, his unit's missions, his available training resources, and other pertinent factors which affect his training" (TMD-1).



Methods. The ARTEP consists of publications currently being developed and validated that will replace applicable Army Training Programs (ATP) and Army Training Tests (ATT). Each ARTEP provides a systematic listing of training and evaluation outlines which contain minimum collective training objectives pertaining to specific missions together with guidance on how to use this information. ARTEPs are used by evaluators to determine the degree to which training objectives have been achieved.

In using the ARTEP for evaluation, a chief evaluator and his staff first develop an evaluation plan. This requires selecting and organizing the ARTEP training and evaluation (T&E) outlines into a logical testing sequence. The T&E outlines consist of (1) statements of the general conditions under which a mission is performed, (2) the primary training/evaluation standard upon which the element performing the mission will be evaluated, and (3) the performance-oriented training objectives which describe the tasks, conditions, and training/evaluation standards for the mission. The T&E outline also specifies the suggested support requirements for conducting an evaluation of the mission. Three considerations are specified as being important in the development of the evaluation plan: (1) the resource requirements for conducting the evaluation, (2) the type and number of units to be evaluated, and (3) the minimum requirements needed to conduct a valid evaluation.

The ARTEP emphasizes the importance of selecting chief evaluators who are highly qualified. It is suggested that the chief evaluator be an individual who (1) has successfully commanded a unit similar to the one he will evaluate, (2) has the confidence of the senior commander who directed that the evaluation be conducted, (3) has personal knowledge of the ARTEP evaluation standards; and (4) is selected from outside the unit being tested. In addition to the importance of the initial selection of highly qualified evaluators, the ARTEP stresses the importance of evaluators receiving adequate training prior to conducting an ARTEP evaluation. This training should include knowledge of the missions to be evaluated and the training/evaluation standards to be met.

INDIVIDUAL EVALUATION. Concepts and practices underlying the assessment of individual proficiency are also being changed within the Army. The Military Occupational Specialty (MOS) tests which have been the primary means of assessing individual proficiency will gradually be supplanted by Skill Qualification Tests (SQT). The current MOS tests are largely performance-oriented paper and pencil tests, whereas the SQTs will feature a high degree of hands-on testing to determine an individual's ability to perform critical tasks essential to success in combat.

The recently formed Individual Training and Evaluation Division (ITED) at TRADOC, Fort Eustis, VA, will be responsible for the development of SQTs for Army-wide use. These tests will be structured to evaluate a soldier's proficiency in his MOS and duty position at the

TAEG Report No. 39

current and next higher skill level. Much developmental work remains to be done on these tests, however, and they must be field-tested and validated before replacing MOS tests.

### SECTION III

#### SUMMARY AND DISCUSSION

##### SUMMARY

Variability exists among the services in the emphasis given to the evaluation of training, as well as in philosophies, concepts, and practices of evaluation. Some of this can be attributed to differences in missions of the services which have brought about unique concepts of the role of training and its attendant evaluation. Other variability must be attributed to an apparent failure in some instances to recognize the need for sustained efforts to obtain objective data regarding training effects which can be used to control and direct training towards desired ends. Command support and resources made available for implementing and sustaining evaluation programs are considerably less than is needed for effective quality control of training.

The Air Force training evaluation program differs significantly from those of the other services. The Air Force has established a formal quality control of training system to maintain surveillance over training. Control of training is achieved through periodically conducted evaluations of all courses and graduates of these courses. The intent is to insure that technical training courses are producing individuals with the skills required for effective job performance. Training evaluations are conducted by individuals who function independently of the training staff. Training deficiencies and action programs to eliminate them are monitored by Headquarters ATC. Course status is reported directly to the Commanding General of ATC.

Navy and Marine Corps training evaluation programs apparently receive less emphasis and less command support than those of the Air Force. While the need for evaluation data is recognized, its collection and use is delegated to the activity which conducts the training. Systematic evaluation programs are not in effect and the command emphasis seems to be one of providing support in the form of general guidance documents and some limited tools (e.g., checklists) for use by the "evaluators." Training for individuals who must serve evaluation roles is less than adequate. Within the Navy, interest in and requirements for evaluation programs are increasing, however. Currently, concerted attempts are being made to develop more adequate procedures for obtaining data from Fleet units reflecting the ability of graduates to perform required jobs. These data, fed back into the training system, should be useful for training improvement.

The Army concept of training effectiveness is different than that of the other services. Training is evaluated in terms of its contribution to combat readiness. Deficiencies in readiness are corrected by

more training. There is no deliberate emphasis placed on assessing the effectiveness of particular courses.

Currently, the Air Force and Marine Corps rely heavily on questionnaires for obtaining data for use in evaluating the effectiveness of training. The Navy plans to use this technique to a much greater extent in the future. Data obtained from questionnaires are not optimum for evaluating training effectiveness. Many factors affect the interpretation, validity, and usefulness of the information that is obtained (e.g., number of items, style, clarity, inclusiveness). At best, obtained data represent subjective opinions. As such, they do not directly reveal the graduate's actual ability to perform job tasks. For the improvement of training, actual performance data are required and objective performance testing programs should be initiated. Hopefully, these programs would also include provisions for validating questionnaire results against job performance criteria. The Army's lead in performance-oriented testing is worthy of emulation by the other services.

#### DISCUSSION OF EVALUATION PROGRAMS

Some apparent deficiencies in training evaluation programs were noted by the project staff during this review effort. These are discussed here separately for each of the services.

**U.S. AIR FORCE.** Air Force personnel (at Lowry Air Force Base) identified a number of features of the Air Force Graduate Evaluation Program that, in their opinions, if changed, would strengthen the program. Opinions were voiced in three areas: procedures, personnel, and resources.

Procedures. Air Force personnel have suggested that the current field evaluation program does not meet all current evaluation needs. For example, in some cases the need is to maintain routine quality control of training. In others, it is to intensively evaluate new or revised courses, or courses which are having other than routine problems. A more differentially applicable field evaluation program employing different evaluation methodologies for different problems is desired. It is understood that this need has been recognized by ATC and that ATC plans to provide separate field evaluation programs for (1) routine quality control, (2) intensive evaluation of new or revised courses, and (3) follow-up evaluation of course changes made as a result of previous evaluations.

The validity of the information obtained by questionnaires has also been questioned by Air Force personnel. They note that there are sources of error in the questionnaires which should be investigated and eliminated. They suggested that a continuing comparison be made at each training center of data obtained using alternate questionnaire form layouts, item content differences, and answer sheet formats. Data obtained by



### TAEG Report No. 39

questionnaire regarding graduates who have worked on the job from 3 to 6 months may also be contaminated. These graduates receive varying amounts of on-the-job, field detachment, or other training during this period. This additional training may contaminate evaluation of the adequacy of their technical school training. No empirical data are available to assist in determining to what obtained results may be reasonably attributed.

A need for development of a computerized routine for selecting samples was also expressed. This would facilitate identification of individuals from/about whom performance information would be collected for course evaluation. An additional problem concerns the type of sample needed for evaluation. Identification needs to be made of the extent to which the sample of graduates evaluated is representative of trainees to be expected in the future. This is necessary to insure that changes made in training as a result of evaluation will be relevant for future trainees. A need for a better data analysis program was also noted. Analysis of data collected from field evaluation questionnaires has been limited to simple frequency counts and percentages. Information is not provided about relationships among variables such as aptitude, educational level, and course performance.

Personnel. At present, individuals who are working as evaluators have received little, if any, formal training in evaluation. A program of inservice training for evaluators to increase their capability and skills for conducting training evaluation was cited as a need. It was also noted that the accuracy of the data regarding graduate performance which is submitted by field supervisors is unknown. Information is needed regarding the motivation and interest of supervisors in evaluating graduates.

Resources. Two major concerns, time and money, were addressed in the area of resources. Under the present graduate evaluation program, a little over 1 year, from start to finish, is required to evaluate a course. Course staff, training requirements, and trainee characteristics change over time and this long time period decreases the relevancy of the evaluative feedback. To improve the relevancy and usefulness of the feedback data, time required for the evaluation process should be shortened. Also, there are limited funds for field evaluation visits. This places the major emphasis of data collection on the questionnaire method with its more limited response capability and inherent weaknesses.

U.S. NAVY. Navy training evaluation is not as systematized as that of the Air Force nor is there a cadre of individuals whose full-time duties involve only evaluation. Comments on aspects of Navy training evaluation are offered below.

Training Appraisal Plans. Basic policies for training evaluation are established and promulgated at command levels. Training activities develop

and implement compliance plans and programs in accordance with Command instructions. The TAEG project team obtained a number of Training Appraisal Plans (TAP) from CNET (N-34) for review. These TAPs were prepared by individual training organizations in response to the requirements of CNET Instruction 1540.3, Appraisal and Improvement of Training. The plans were reviewed to assess the ability of training units to comply with the Instruction.

The TAPS varied considerably in length, quality, thoroughness, and apparent understanding of the requirements in terms conducive to satisfying them. Most of the plans specifically addressed the topics required by the instruction but typically did not describe how evaluations would proceed. Also, evaluation standards were not adequately addressed. Some TAPs referenced CNET or CNTECHTRA instructions regarding standards (e.g., measurement of student achievement will be "as good as" specified in CNET Instruction 1540.2).

Testing programs for obtaining evaluation data were noticeably lacking. Provisions for comparative testing (e.g., alternate learning strategies, alternate configurations of resources) to determine if training could be done more effectively and/or efficiently were absent. Many schools indicated in their TAPs that they have evaluation divisions but did not indicate who staffs them or if it is full-time or collateral duty. As a point of interest, a recent TAEG survey in which data were obtained on 435 "A" and "C" courses shows that only 45 percent of these give comprehensive end-of-course examinations. Increased emphasis on testing to determine what the students learn in courses is needed.

Internal evaluation was addressed more fully in most TAPs than was external evaluation. However, some schools described rather elaborate systems for obtaining feedback of evaluative information from the Fleet. Others included no provision for formal external evaluation. The TAEG survey mentioned above also revealed that only 20 percent of the courses collect feedback information (by questionnaires of unknown description) from graduates and their supervisors. Seventy percent rely on informal feedback from Fleet personnel. The plans were also weak in describing how evaluation information would be used to improve training. In the area of instructor evaluation, for example, most plans specify that instructors will be evaluated according to guidance promulgated by CNTECHTRA Instruction 1540.12. Some plans were more specific and did indicate that specific programs of inservice training would be conducted to correct deficiencies in instructor performance.

<sup>1</sup> Unpublished data collected under TAEG Work Assignment W1065, May 1976.

CNET Instruction 1540.3A, Appraisal and Improvement of Training, which is currently in preparation, will cancel CNET Instruction 1540.3. The essential provisions of 1540.3A are described in section II of this report. Certain of these provisions have important implications for the CNET quality control of Navy training. For example, the responsibility for evaluation still rests ultimately in the hands of the school personnel who must also conduct training rather than being vested in an independent evaluation group. The stated purpose of the instruction is to "provide information for standardizing training appraisal." The assumption seems to be that this can be done by obtaining feedback data from the Fleet regarding how well course graduates perform, and the major emphasis of the instruction is concerned with training appraisal surveys. While useful information can be gathered for improving training, it is difficult to see how such information will assist in "standardizing training appraisal." The reliance on subjective opinion data rather than on objective data reflecting graduates' ability to perform job tasks is also less than desirable. The instructions suggest that development of single evaluation plans that can be used by all CNET activities is not practical due to the diversity of training within the NAVEDTRACOM. This diversity of training situations obviously also means that there are likewise many diverse evaluation situations. Thus, selection of techniques for assessing training effectiveness should be based on consideration for the specific elements of a given situation. This means that a diversity of potential assessment techniques should be examined to determine their applicability to particular cases. Indiscriminate use of single techniques (e.g., questionnaires) for obtaining "effectiveness" information is not recommended. Volume II of this study discusses the problems of conducting training effectiveness assessments. It also provides information concerning a variety of techniques for obtaining objective job performance data for use in training evaluation.

The proposed CNET Instruction 1540.3A does not delegate final approval authority for conducting training appraisal surveys to CNET. At present, it is planned that this authority will be assigned to the Chief of Naval Personnel. CNET, however, is responsible for insuring that the training system meets the goals for which the NAVEDTRACOM was established. Thus, it would seem that final approval authority for conducting surveys to obtain information for evaluating training should be vested in CNET.

Curriculum and Instructional Standards Offices. A requirement for the creation of Curriculum and Instructional Standards Offices (CISO) has been established by CNET Instruction 1540.6 (July 1975). Its provisions and the functions to be served by these CISOs have already been described in section II of this report. Several provisions of that instruction are of interest. For example, the instruction clearly recognizes the need for effective controls, or checks, over the training process and attempts to introduce them within the current training organization.



Provisions for the administrative separation of the evaluation branch from the instructional department are noteworthy. Hopefully, this will minimize biasing evaluation and assist in the production of meaningful evaluative data. The instruction states, however, that "manpower allocations to training activities will not be increased in that the majority of the functions of the CIS offices or departments are already being performed by training activity personnel." But the duties described are quite technical and also very comprehensive. Thus, without increased manpower, difficulties in fully complying with the instruction can be predicted. Also, without manpower increases, it is difficult to see how training and evaluation functions can be effectively separated. At present, CIS offices are now being formed within the Naval technical training organization and current indications are that there will be organizational variability in the manner in which assigned duties will be accomplished.

CNTECHTRA AIO Manual. The evaluation programs and procedures described in the CNTECHTRA AIO Manual appear to be basically sound. But, unfortunately, there appears to be no formal requirement that the procedures be applied, and it is believed that internal evaluation checklists are not widely used to evaluate Navy courses. Some school personnel have stated that they do not have the time to conduct the recommended evaluations. One of the more significant problems of the manual concerns the issue of standards, or criteria, by which to judge the quality of aspects of a course. Many of the internal evaluation checklist items are to be assigned a rating of "adequate," "generally adequate," "inadequate," etc., but, the definition of what constitutes "adequate," etc., is left to the subjective judgment of the evaluator. For example, one of the review elements in the instructional staff checklist is, "provide for ample student/instructor/learning supervisor interaction." There is, however, no explanation of what constitutes adequate or inadequate "ample student/instructor interaction." Without more specific, objective definitions of these evaluative standards, ratings will more frequently reflect the biases of evaluators rather than the true condition of the course elements.

U.S. MARINE CORPS. Marine Corps documentation describing evaluation needs, practices to follow, and uses of data is comprehensive and seems to address adequately those areas which should be considered for evaluation. But, evaluation, as in the Navy, is vested largely in the training organization itself rather than within a separately-constituted and impartial evaluation group. The ability of assigned personnel to fully understand, implement, and execute evaluation programs is not known and the value of the programs is uncertain. A high degree of reliance is placed on student comments/ratings of courses (IRFs) and instructor evaluations (AIRs) for improving and/or maintaining course quality. While these sources may provide important clues regarding deficiencies in training, they should certainly be supplemented with more objective data regarding what students actually learn in the course and how



TAEG Report No. 39

different learning experiences affect their achievement. The use of questionnaires for obtaining data regarding graduate performance in the field should be validated against data obtained through job performance testing.

U.S. ARMY. The Army evaluation program as it is now conceived focuses on the assessment of individual and unit proficiency. The intent in both instances is to determine readiness, or ability, to perform critical missions. When assessments/evaluations are made, no deliberate attempt is made to attribute observed proficiency to institutional training. Rather, the attempt is to discover if additional training is needed. This concept and its concomitant practices presumably meet the Army's needs and provide the evaluative information that the Army feels is necessary for insuring, or progressing towards, full combat readiness. The program is not intended to provide information about the effects of particular training courses, and quality control of the training system does not appear to be an issue. Thus, no attempt is made to determine what or how much formal training contributes to an individual soldier's skills.

REFERENCES

U.S. AIR FORCE

Air Training Command. Technical Training: Training Evaluation and Course Reviews. ATC Regulation 52-1. 1 November 1974. Headquarters, Randolph Air Force Base, TX 78148.

Department of the Air Force. Training: Field Evaluation of Formal School Graduates. AF Regulation 50-38. 30 April 1975. Headquarters, U.S. Air Force, Washington, D.C. 20330.

Air Training Command. Technical Training: Student Critique Program. ATC Regulation 52-29. 23 January 1975. Headquarters, Air Training Command, Randolph AFB, TX 78148.

Air Training Command. Technical Training. Student Measurement. ATC Regulation 52-3. 8 January 1973. Headquarters, Air Training Command, Randolph AFB, TX 78148.

Department of the Air Force. Training: Handbook for Designers of Instructional Systems (Evaluation). Volume 5, Air Force Pamphlet 50-58. 15 July 1973. Headquarters, USAF, Washington, D.C. 20330.

USAF School of Applied Aerospace Sciences (ATC). Technical Training: Training Evaluation and Course Reviews. School Supplement 1 to ATC Regulation 52-1. 31 March 1975. Headquarters, USAF School of Applied Aerospace Sciences, Lowry AFB, CO 80230.

U.S. NAVY

Naval Education and Training Command. Interservice Procedures for Instructional Systems Development. NAVEDTRA 106A. 1975. Naval Education and Training Command, Pensacola, FL 32508.

Chief of Naval Technical Training. Procedures for the Planning, Design, Development, and Management of Navy Technical Training Courses. CNTT AIO (Manual). April 1976. Chief of Naval Technical Training, Millington, TN 38054.

Chief of Naval Education and Training. Training Appraisal Subsystem; Establishment of Curriculum and Instructional Standards (CIS) Offices/Departments. CNET Instruction 1540.6. 11 July 1975. Chief of Naval Education and Training, Pensacola, FL 32508.

Chief of Naval Education and Training. Measurement of Student Achievement. CNET Instruction 1540.2. 13 August 1973. Chief of Naval Education and Training, Pensacola, FL 32508.

TAEG Report No. 39

REFERENCES (continued)

Chief of Naval Education and Training. Appraisal and Improvement of Training. CNET Instruction 1540.3. 23 October 1973. Chief of Naval Education and Training, Pensacola, FL 32508.

Chief of Naval Education and Training. Appraisal and Improvement of Training. CNET Instruction 1540.3A (Draft). (no date). Chief of Naval Education and Training, Pensacola, FL 32508.

Chief of Naval Technical Training. Evaluation of Instruction. CNTECHTRA Instruction 1540.12. 24 January 1973. Chief of Naval Technical Training, Millington, TN 38054.

Walker, C.L. (LCDR). The Personnel and Training Evaluation Program: A Working Program for Improving the Efficiency and Effectiveness of Fleet Ballistic Missile Weapons Systems Training. Part I, Scope and Achievement. Unpublished Paper. (no date). Central Test Site for the Personnel and Training Evaluation Program, Dam Neck, VA

Braun, F.B. The Personnel and Training Evaluation Program: A Working Program for Improving the Efficiency and Effectiveness of Fleet Ballistic Missile Weapons System Training. Part II, Program Development. Unpublished Paper. (no date). Data Design Laboratories, Norfolk, VA.

Central Test Site for the Personnel and Training Evaluation Program. Fleet Ballistic Missile Weapons System Training Program: Personnel and Training Evaluation Program. General Discription. 1 August 1974. Central Test Site for PTEP, Dam Neck, VA.

Dyer, F.N., Ryan, L.E., and Mew, Dorothy V. A Method for Obtaining Post Formal Training Feedback: Development and Validation. TAEG Report No. 19. May 1975. Training Analysis and Evaluation Group, Orlando, FL 32813.

Dyer, F.N., Mew, Dorothy V., and Ryan, L.E. Procedures for Questionnaire Development and Use in Navy Training Feedback. TAEG Report No. 20. October 1975. Training Analysis and Evaluation Group, Orlando, FL 32813.

U.S. ARMY

U.S. Army Infantry School. "Training Management: An Overview." Training Management Digest, TC-21-5-1, Number One. April 1973. Ft. Benning, GA.

U.S. Army Infantry School. "Performance - Oriented Training." Training Management Digest, TC-21-5-2, Number Two. June 1974. Ft. Benning, GA.

TAEG Report No. 39

REFERENCES (continued)

Headquarters, United States Army Training and Doctrine Command. Analyzing Training Effectiveness, TRADOC PAM 71-8. Ft. Monroe, VA.

Headquarters, Department of the Army. How to Prepare and Conduct Military Training. FM 21-6. November 1975. Washington, DC.

U.S. MARINE CORPS

GENTRAS Functional Description. Extracted from a document prepared in support of Task G of Contract M00027-70-C-0100, Project Serial Number CF 0189001 for Headquarters U. S. Marine Corps, Washington, DC.

U.S. Marine Corps Command and Staff College. Program of Instruction for Academic Year 1975-1976. Marine Corps Command and Staff College, Quantico, VA.

U.S. Marine Corps. Design of Courses of Instruction. MCO P1510.23A. 7 November 1972. Headquarters, U.S. Marine Corps, Washington, DC.

U.S. Marine Corps Development and Education Command. Academic Regulations, ECO P5000.IJ. 25 February 1975. U.S. Marine Corps Development and Education Command, Quantico, VA.

U.S. Marine Corps. Unit Level Training Management. MCO P1510.26. 4 May 1971. Headquarters U.S. Marine Corps, Washington, D.C.



TAEG Report No. 39

DISTRIBUTION LIST

CNET (00A, N-5 (5 copies))  
CNET Support (00, 01A)  
CNTECHTRA (0161, Dr. Kerr (5 copies); Library)  
CNATRA (F. Schufletowski)  
CNAVRES (Code 02)  
COMTRALANT  
COMTRALANT (Educational Advisor)  
COMTRAPAC  
CO NAVEDTRASUPPCEN NORVA  
CO NAVEDTRASUPPCENPAC (5 copies)